Non-Standard Asymptotics in High Dimensions: Manski’s Maximum Score Estimator Revisited

Manski’s celebrated maximum score estimator for the censored response linear model has been the focus of much investigation in both the econometrics and statistics literatures but its behavior under growing dimension scenarios still largely remains unknown. Two different cases: $p$ grows with $n$ but at a slow rate, i.e., $\frac{p}{n} \rightarrow 0$; and $p > n$ (fast growth) are considered. By relating Manski’s score estimation to empirical risk minimization in a classification problem, we show that under a soft margin condition (due to Tsybakov) involving a smoothness parameter $\alpha > 0$, the rate of the score estimator in the slow regime is essentially $(\frac{p}{n} \log n)^{\frac{\alpha}{2}}$, while, in the fast regime, the $l_0$ penalized score estimator essentially attains the rate $(s_0 \log p \log n/n)^{\frac{\alpha}{2}}$. For the most interesting regime, $\alpha = 1$, the rates of Manski’s estimator are therefore $(\frac{p}{n} \log n)^{1/3}$ and $(s_0 \log p \log n/n)^{1/3}$ in the slow and fast growth scenarios respectively, which can be viewed as high-dimensional analogues of cube-root asymptotics (Kim and Pollard, 1990, Annals of Statistics). We also establish upper and lower bounds for the minimax $L_2$ error in Manski’s model that differ by a logarithmic factor. Finally, we provide computational recipes for the maximum score estimator in growing dimensions that show promising results. Joint work with Debarghya Mukherjee and Ya’acov Ritov.

Refreshments will be served following the seminar in 1181 Comstock Hall.